A multilevel stochastic collocation method for SPDEs

A multilevel stochastic collocation method for SPDEs

Max Gunzburger[*], Peter Jantsch[†], Aretha Teckentrup[**] and Clayton Webster[‡]

[*]*Florida State University*
[†]*University of Tennessee*
[**]*Florida State University; current address: University of Warwick*
[‡]*Oak Ridge National Laboratory*

Abstract. We present a multilevel stochastic collocation method that, as do multilevel Monte Carlo methods, uses a hierarchy of spatial approximations to reduce the overall computational complexity when solving partial differential equations with random inputs. For approximation in parameter space, a hierarchy of multi-dimensional interpolants of increasing fidelity are used. Rigorous convergence and computational cost estimates for the new multilevel stochastic collocation method are derived and used to demonstrate its advantages compared to standard single-level stochastic collocation approximations as well as multilevel Monte Carlo methods.

Keywords: multilevel methods, stochastic collocation, SPDEs, uncertainty quantification, finite element methods
PACS: 02.50.Fz, 05.40.-a, 02.30.Jr, 02.70.Tt

INTRODUCTION

Consider the elliptic PDE problem with finite dimensional noise

$$\begin{cases} -\nabla\cdot(a(\boldsymbol{y},\mathbf{x})\nabla u(\boldsymbol{y},\mathbf{x})) &= f(\boldsymbol{y},\mathbf{x}) \quad \text{in } \Gamma\times D \\ u(\boldsymbol{y},\mathbf{x}) &= 0 \quad \text{on } \Gamma\times\partial D, \end{cases} \tag{1}$$

where D denotes a bounded spatial domain having boundary ∂D, $\boldsymbol{y}=[y_1(\omega),\dots y_N(\omega)]\in\Gamma$ denotes a vector of bounded independent random variables having a joint probability density function $\rho(\boldsymbol{y})$, and, without loss of generality, we set $\Gamma=[-1,1]^N$. The random variables could arise from a truncation of an infinite representation, e.g., a Karhunen-Loève expansion, of a correlated random field or could be independent random physical parameters. We assume throughout that the problem (1) is subject to a finite element spatial discretization; for any $\boldsymbol{y}\in\Gamma$, we let $u^h(\boldsymbol{y},\mathbf{x})$ denote the corresponding finite element approximation. The formulation of the multilevel method does not depend on the specific spatial discretization scheme used and the results readily hold for other choices [1].

The most commonly used approach towards obtaining statistical information about the solution u of (1) is to apply a Monte Carlo method wherein M points $\{\boldsymbol{y}_m\}_{m=1}^M$ chosen at random within Γ, then the finite element system corresponding to (1) is solved M times, once for each $\boldsymbol{y}_m$, to obtain the M solutions $\{u^h(\boldsymbol{y}_m,\mathbf{x})\}_{m=1}^M$, and then, e.g., if one is interested in the expectation of u and uniform sampling is used, one computes the ensemble average $\frac{1}{M}\sum_{m=1}^M u^h(\boldsymbol{y}_m,\mathbf{x})\rho(\boldsymbol{y}_m)$.

To mitigate the known slow convergence of Monte Carlo methods, *multilevel Monte Carlo* methods have been recently developed; see, e.g., [2, 3, 4, 5, 6, 7, 8]. A hierarchy of nested finite element subspaces $V_{h_0}\subset V_{h_1}\subset\cdots\subset V_{h_K}\subset H_0^1(D)$ is chosen, where the grid size parameters h_k, $k=0,\dots,K$, are decreasing with k, e.g., $h_k\propto 1/2^k$. We can write u_{h_K} as the simple telescoping sum

$$u_{h_K}=u_{h_0}+\sum_{k=1}^{K}(u_{h_k}-u_{h_{k-1}}). \tag{2}$$

In general, the difference $\Delta_k=(u_{h_k}-u_{h_{k-1}})$ between successive finite element approximation becomes smaller as k increases. Thus, when computing ensemble averages, one can sample u_{h_0} and Δ_k for small k many times and sample Δ_k for larger k a decreasing number of times without affecting the overall accuracy of the approximation. Because relatively large number of samples are done for coarse grid (and therefore cheaply obtained) finite element approximations and a relatively small number of samples are done for fine grid (and therefore expensively obtained) finite element approximations, the overall strategy results in considerable savings over Monte Carlo methods that simply sample u_{h_K}.

Stochastic collocation methods [9, 10, 11] are another means of improving on MC methods. Here, deterministic sample points are used; of particular interest for this purpose are *sparse grid* points such as those used for Smolyak quadrature. If the solution has smooth dependence on the random variables $\boldsymbol{y}$, as is the case for the problem (1), then it has been shown that stochastic collocation methods converge much faster than do Monte Carlo methods. For details about tochastic collocation methods and sparse grids, see [9, 10, 11, 12].

In this work, we study the convergence and complexity behaviors of a multilevel stochastic collocation method which employs the same philosophy as do multilevel Monte Carlo methods. Hierarchies of finite element spaces and sparse grids are defined and then coarse (fine) sparse grids in $\boldsymbol{y}$ are used for finite element computations on fine (coarse) spatial grids. We conclude that the multilevel stochastic collocation method converges faster for the same cost than do ordinary stochastic methods and multilevel Monte Carlo methods. We remark that the multilevel stochastic collocation we study is not specific to the model problem (1); it can be applied also to other partial differential equations, including higher order and nonlinear ones, and to other types of boundary conditions.

HIERARCHICAL MULTILEVEL STOCHASTIC COLLOCATION METHODS

We begin by recalling that standard stochastic collocation methods generally build an approximation of the solution u by evaluating a spatial approximation $u_h(\boldsymbol{y},\cdot) \in V_h$ at a given set of points $\{\boldsymbol{y}_m\}_{m=1}^M$ in Γ, where $V_h \subset H_0^1(D)$ is a finite-dimensional subspace. In other words, we compute $\{u_h(\boldsymbol{y}_m,\cdot)\}_{m=1}^M$. Then, given a basis $\{\phi_m(\boldsymbol{y})\}_{m=1}^M$ for the space $\mathscr{P}_M = \operatorname{span}\{\phi_m(\boldsymbol{y})\}_{m=1}^M \subset L_\rho^2(\Gamma)$, we use those samples to construct the fully discrete approximation given by the interpolant

$$u_{M,h}^{(\mathrm{SC})}(\boldsymbol{y},\mathbf{x}) = \mathscr{I}_M[u_h](\boldsymbol{y},\mathbf{x}) = \sum_{m=1}^{M} c_m(\mathbf{x})\,\phi_m(\boldsymbol{y}), \tag{3}$$

where the coefficients $c_m(\mathbf{x})$ are determined by the interpolating conditions $\mathscr{I}_M[u_h](\boldsymbol{y}_m,\mathbf{x}) = u_h(\boldsymbol{y}_m,\mathbf{x})$ for $m = 1,\ldots,M$.

To obtain a good approximation with stochastic collocation methods, it is necessary in general to use accurate spatial approximations u_h and a large number M of collocation points, i.e., the computation of $u_h(\boldsymbol{y}_m,\cdot)$ for $m = 1,\ldots,M$, so that, in practice, the cost can grow quickly with increasing N. Therefore, to reduce the overall cost, we consider a multilevel version of stochastic collocation methods that combines different levels of fidelity of both the spatial and parameter approximations.

For *spatial approximation*, we use hierarchical family of finite element discretizations based on a hierarchy of nested finite element spaces $\{V_{h_k}\}_{k=0}^K$, where each $V_{h_k} \subset H_0^1(D)$ consists of continuous, piecewise polynomial functions with respect to the corresponding grid. We assume that there exist positive constants α and C_s, independent of h_k, such that for all $k \in \mathbb{N}_0$, $\|u - u_{h_k}\|_{L_\rho^2(\Gamma;H_0^1(D))} \le C_s\, h_k^\alpha$.

For *stochastic approximation*, we use a sequence of interpolation operators $\left\{\mathscr{I}_{M_k}\right\}_{k=0}^\infty$ using M_k points and assume that $\|v - \mathscr{I}_{M_k} v\|_{L_\rho^2(\Gamma;H_0^1(D))} \le C_I\,\sigma_k\,\zeta(v)$ for some function $\zeta : \Lambda(\Gamma;H_0^1(D)) \to \mathbb{R}$ and a decreasing sequence σ_k that admit the estimates $\zeta(u_{h_k}) \le C_\zeta\, h_0^\beta$ and $\zeta(u_{h_{k+1}} - u_{h_k}) \le C_\zeta\, h_{k+1}^\beta$. Note that it is possible and even desirable that, for any $k = 0,\ldots,K$, $M_{k+1} = M_k$. Thus, although the spatial approximation improves with increasing k, i.e., $h_{k+1} < h_k$, we allow for the parameter space approximation for the index $k+1$ remaining the same as that for k. Also note that the assumptions hold with $\sigma_k = M_k^{-\mu}$ for global Lagrange interpolation using generalized sparse grids.

The assumptions made imply that as k increases, less accurate interpolation operators are needed in order to estimate $u_{h_k} - u_{h_{k-1}}$ to achieve a required accuracy. Therefore, we define the multilevel stochastic collaction approximation as

$$u_K^{(\mathrm{MLSC})} := \sum_{k=0}^{K} \mathscr{I}_{M_{K-k}}[u_{h_k} - u_{h_{k-1}}] = \sum_{k=0}^{K} \left(u_{M_{K-k},h_k}^{(\mathrm{SC})} - u_{M_{K-k},h_{k-1}}^{(\mathrm{SC})}\right). \tag{4}$$

Rather than simply interpolating u_{h_K}, this approximation uses different levels of interpolation on each difference $u_{h_k} - u_{h_{k-1}}$ of finite element approximations. To preserve convergence, the estimator uses the most accurate interpolation operator $\mathscr{I}_{M_K}$ on the coarsest spatial approximation u_{h_0} and the least accurate interpolation operator $\mathscr{I}_{M_0}$ on the finest spatial approximation $u_{h_K} - u_{h_{K-1}}$.

Convergence and cost analyses

The triangle inequality is used to split the error into the sum of a spatial discretization error and a stochastic interpolation error so that each is estimated separately. Choosing interpolation operators such that $\sigma_{K-k} \leq C_s \left((K+1) C_I C_\zeta\right)^{-1} h_K^\alpha h_k^{-\beta}$, we then obtain that $\|u - u_K^{(\mathrm{MLSC})}\|_{L^2_\rho(\Gamma;H^1_0(D))} \leq 2 C_s h_K^\alpha$.

To estimate the computational cost of the MLSC method, we consider the *ε-cost of the estimator*, denoted here by $C_\varepsilon^{\mathrm{MLSC}}$, which is the computational cost required to achieve a desired accuracy ε. We assume $\sigma_k = M_k^{-\mu}$ for some $\mu > 0$. We also assume that there exist positive constants γ and C_c, independent of h_k, such that $C_k \leq C_c h_k^{-\gamma}$, where C_k denotes the cost of computing $u_{h_k} - u_{h_{k-1}}$ at a sample point. For example, if an optimal linear solver is used to solve the finite element equations for u_{h_k}, this assumption holds with $\gamma \approx d$. Note that the constant C_c will in general depend on the refinement ratio η. The total computational cost for obtaining the multilevel stochastic collocation approximation (4) is then given by $C^{(\mathrm{MLSC})} = \sum_{k=0}^K M_{K-k} C_k$. We then can prove the following estimate: assume that $\alpha \geq \min(\beta, \mu\gamma)$; then, for any $\varepsilon < \exp[-1]$, there exists an integer K such that $\|u - u_K^{(\mathrm{ML})}\|_{L^2_\rho(\Gamma;H^1_0(D))} \leq \varepsilon$ and

$$C_\varepsilon^{(\mathrm{MLSC})} \lesssim \begin{cases} \varepsilon^{-\frac{1}{\mu}}, & \text{if } \beta > \mu\gamma \\ \varepsilon^{-\frac{1}{\mu}} |\log \varepsilon|^{1+\frac{1}{\mu}} & \text{if } \beta = \mu\gamma \\ \varepsilon^{-\frac{1}{\mu} - \frac{\gamma\mu - \beta}{\alpha\mu}} & \text{if } \beta < \mu\gamma. \end{cases} \tag{5}$$

For the standard, single-level stochastic collocation method, we have $C_\varepsilon^{(\mathrm{SC})} \lesssim \varepsilon^{-\frac{1}{\mu} - \frac{\gamma}{\alpha}}$. A comparison with (5) clearly shows the superiority of the multilevel method in all cases.

The multilevel stochastic collocation scheme described above has been applied to several sparse grid collocation methods and for which we have verified the truth of all the assumptions made. Here, we content ourselves to simply say that the stochastic collocation and multilevel stochastic collocation methods both outperform Monte Carlo and multilevel Monte Carlo methods and that multilevel stochastic collocation methods methods outperform stochastic collocation methods. Furthermore, the gaps in performance increase dramatically as the dimension N of the parameter space increases.

REFERENCES

1. H. Harbrecht, M. Peters, and M. Siebenmorgen, "On multilevel quadrature for elliptic stochastic partial differential equations," in *Sparse Grids and Applications*, Springer, 2013, pp. 161–179.
2. M. Giles, *Oper. Res.* **56**, 607–617 (2008).
3. A. Barth, C. Schwab, and N. Zollinger, *Numer. Math.* **119**, 123–161 (2011).
4. J. Charrier, R. Scheichl, and A. Teckentrup, *SIAM J. Numer. Anal.* **51**, 322–352 (2013).
5. S. Dereich, and F. Heidenreich, *Stoch. Proc. Appl.* **121**, 1565–1587 (2011).
6. C. Gittelson, J. Könnö, C. Schwab, and R. Stenberg, *Numer. Math.* **125**, 347–386 (2013).
7. M. Giles, and C. Reisinger, *SIFIN* **1**, 575–592 (2012).
8. H. Hoel, E. von Schwerin, A. Szepessy, and R. Tempone, "Adaptive multilevel Monte Carlo simulation," in *Numerical Analysis of Multiscale Computations*, edited by B. Engquist, O. Runborg, and Y.-H. R. Tsai, Springer, 2012, vol. 82 of *Lect. Notes Comp. Sci.*, pp. 217–234.
9. I. Babuška, F. Nobile, and R. Tempone, *SIAM J. Numer. Anal.* **45**, 1005–1034 (2007).
10. F. Nobile, R. Tempone, and C. Webster, *SIAM J. Numer. Anal.* **46**, 2309–2345 (2008).
11. F. Nobile, R. Tempone, and C. Webster, *SIAM J. Numer. Anal.* **46**, 2411–2442 (2008).
12. J. Beck, F. Nobile, L. Tamellini, and R. Tempone, *Lect. Notes Comp. Sci.* **76**, 43–62 (2011).